

## Away from the edge: SAD phasing from the sulfur anomalous signal measured in-house with chromium radiation

**Cheng Yang, J. W. Pflugrath,  
D. A. Courville, C. N. Stence and  
Joseph D. Ferrara\***

Rigaku/MSD Inc., 9009 New Trails Drive,  
The Woodlands, TX 7738, USA

Correspondence e-mail: jdf@rigakumsc.com

Received 29 April 2003  
Accepted 19 August 2003

Anomalous scattering with soft X-ray radiation opens new possibilities in phasing for macromolecular crystallography. Anomalous scattering from S atoms collected on an in-house chromium radiation source ( $\lambda = 2.29 \text{ \AA}$ ) was used to phase the X-ray diffraction data of thaumatin (22 kDa) and trypsin (24 kDa) crystals. The contribution to the anomalous term,  $\Delta f'' = 1.14 \text{ e}^-$ , from sulfur for Cr  $K\alpha$  radiation is doubled compared with that for Cu  $K\alpha$  radiation,  $\Delta f'' = 0.56 \text{ e}^-$ . The direct-methods programs *RANTAN* or *SHELXD* successfully found sulfur positions using data sets with resolution limited to 3.5  $\text{\AA}$ . The statistical phasing program *SHARP* was used to produce the electron-density maps using the sulfur anomalous signal alone at low resolution ( $\sim 3.5 \text{ \AA}$ ). An interpretable electron-density map for each structure was obtained solely from the phases derived from single-wavelength anomalous dispersion (SAD) data obtained using Cr  $K\alpha$  radiation. Much fewer data (that is, lower redundancy) are required for this sulfur SAD phasing procedure compared with the highly redundant data reported in the sulfur SAD phasing procedure with Cu  $K\alpha$  radiation. Cr  $K\alpha$  radiation can also improve the strength of anomalous scattering of many other intrinsic elements in macromolecules, such as calcium, zinc and phosphorus, because of the increased  $\Delta f''$ . Furthermore, the anomalous scattering of selenium is increased substantially from 1.14  $\text{e}^-$  with Cu  $K\alpha$  radiation to 2.28  $\text{e}^-$  with Cr  $K\alpha$  radiation. In order to measure the small Bijvoet differences accurately, several devices were developed for the experiment, including an Osmic Confocal MaxFlux optic optimized for Cr  $K\alpha$  radiation, a helium path and a beam stop. In the cases studied here, radiation damage to the samples and reduction of anomalous signal were observed in some long exposure time data sets. Therefore, an adequate data-collection strategy to maximize the completeness in a short scan range was used in subsequent data collections. The results show that the anomalous signal of S atoms can be collected quickly. Since the absorption of solvent and the loop may no longer be negligible with Cr  $K\alpha$  radiation, the orientation of the crystal and exposure time were taken into account in order to minimize the effects of radiation damage and absorption. This experimental study shows that using Cr  $K\alpha$  radiation from an in-house rotating-anode X-ray generator can provide sufficient phasing power from sulfur anomalous signals to routinely phase protein diffraction data.

### 1. Introduction

Anomalous dispersion methods are widely exploited in macromolecular crystallography. In the past ten years, multi-wavelength anomalous dispersion (MAD; Hendrickson, 1991) has become the most powerful method for determining

protein structures. This method requires a tunable synchrotron X-ray source in order to obtain the largest difference in the  $\Delta f''$  and  $\Delta f'''$  contributions near the absorption edge of an anomalous scattering species. In this approach, two or three data sets are collected at various wavelengths around an absorption edge of the anomalous scatterer present in the crystal. In addition to using intrinsic metal ions such as Zn, Cu, Fe, Mn and Ca as the anomalous scatterers, methionine residues in a protein are often substituted by selenomethionine (SeMet) in order to provide anomalous scatterers. Indeed, MAD phasing eliminates the trial-and-error step of the preparation of heavy-atom derivatives, which are required in the single isomorphous replacement (SIR) and multiple isomorphous replacement (MIR) methods. Additionally, these methods are affected by errors caused by non-isomorphism. In favorable circumstances, the MAD phasing approach can produce accurate phases and excellent experimental electron-density maps. Ideally, the MAD technique enables the structure to be solved using one crystal. However, the MAD technique cannot be easily employed in the home laboratory. It is also known that researchers have failed to prepare some SeMet proteins or to grow crystals of SeMet proteins. At this point in time, the native crystals of most proteins are still grown before any attempt is made to express the protein in the selenomethionine form. The preparation of the SeMet proteins and/or heavy-atom derivatives of native proteins for phase determination remains the rate-limiting step for solving protein structures using the MAD technique after crystallization.

In the present era of high-throughput crystallography, many new techniques have been developed to speed up phase determination. For example, the short cryosoaking procedure has become increasingly popular for rapid heavy-atom derivatization with halides (Dauter & Dauter, 1999; Dauter *et al.*, 2000; Nagem *et al.*, 2001). When the method of single-wavelength anomalous dispersion (SAD) phasing is used with these techniques, it becomes a powerful tool in phasing macromolecular crystal structures (Dauter *et al.*, 2002). Recently, several structures reported have been solved with the SAD phasing technique (Chen *et al.*, 1991; Brodersen *et al.*, 2000; Liu *et al.*, 2000; Wu *et al.*, 2001). SAD phasing from sulfur anomalous scattering is an extremely attractive approach because sulfur is an element that is found in almost all proteins. It does not require substitution of methionine with selenomethionine and the trial-and-error procedure of making heavy-atom derivatives. In principle, a macromolecular structure can be solved with a single data set collected from the native protein when sulfur SAD phasing is employed. In fact, sulfur SAD phasing was successfully used to solve the crambin structure in 1981. This work, performed by Hendrickson & Teeter (1981), led to the structure determination of the 46-residue protein crambin directly from the anomalous scattering of S atoms found in six cysteine residues. The classic work of Wang in the 1980s also advocated the possibility of sulfur SAD phasing (Wang, 1985). More recently, Dauter *et al.* (1999) have shown that small anomalous differences can be used to solve the phase problem for diffraction

data collected from tetragonal hen egg-white lysozyme crystals. They used the anomalous signal from S and Cl atoms measured at a wavelength of 1.54 Å from a synchrotron source with data collected to 1.53 Å resolution. Yang & Pflugrath (2001) and Debreczeni *et al.* (2003) reported that the weak anomalous scattering from S atoms in several protein crystals (thaumatin, trypsin and insulin) can be accurately measured using home sources generating Cu K $\alpha$  radiation (1.54 Å). Such data are sufficient to phase protein data or improve the phases obtained from other methods. The structure of *Escherichia coli* argininosuccinate synthetase has been determined using four different phasing approaches: sulfur SAD, selenium SAD and S/Se-SIRAS data measured with Cu K $\alpha$  radiation (Lemke *et al.*, 2002). Liu *et al.* (2000) determined the crystal structure of the bioluminescent photoprotein obelin directly from its sulfur substructure using synchrotron radiation at a wavelength of 1.74 Å. Borek *et al.* (2002) presented several examples at the 2002 ACA meeting in which protein data was phased using the anomalous signal of sulfur or intrinsic heavy atoms collected at a wavelength of 1.90 Å. Stuhmann and coworkers (Behrens *et al.*, 1998; Stuhmann *et al.*, 1995, 1997) have exploited relatively heroic data collection on single crystals of ribosome, trypsin and bacteriorhodopsin near the sulfur K absorption edge (5.02 Å) where  $\Delta f''$  is maximized. This experiment required a great deal of modification of the hardware setup at a synchrotron beamline and is not practical for routine data collection. However, many earlier efforts to use the weak anomalous signal of S atoms to solve the phase problem were limited to well diffracting crystals of relatively small or sulfur-rich proteins.

Since the beginning of the 1990s, soft X-ray radiation has been used in macromolecular crystallography. Anomalous scattering from soft X-ray radiation opens new possibilities for addressing various problems in macromolecular crystallography (Borek *et al.*, 2002; Behrens *et al.*, 1998; Cianci *et al.*, 2001; Liu *et al.*, 2000; Stuhmann *et al.*, 1995, 1997; Weiss *et al.*, 2001). The wavelength of chromium K $\alpha$  radiation (Cr K $\alpha$ ) from a home-laboratory X-ray source (2.29 Å) is longer than the wavelengths used in most of experiments mentioned above. It appears to be more ideally suited for measuring anomalous signals from weak anomalous scatterers, such as S, Se, Ca and other atoms commonly found in protein crystals. The reason is that the  $\Delta f''$  values for sulfur, calcium and selenium are 1.14, 2.51 and 2.28 e<sup>-</sup>, respectively, at 2.29 Å compared with 0.56, 1.29 and 1.14 e<sup>-</sup>, respectively, at 1.54 Å (*International Tables for X-ray Crystallography*, 1974, Vol. IV, p. 149). The application of Cr K $\alpha$  radiation in X-ray crystallography dates back to the work of Blow (1958). Cr K $\alpha$  radiation was used in the attempt to measure the anomalous signal of an Hg atom in the structure determination of hemoglobin. Unfortunately, the Cr K $\alpha$  data set was not included in the final phase determination because of the large errors in the data. Wang (1985) foresaw the potential of Cr K $\alpha$  radiation for phasing using the sulfur anomalous signal through theoretical calculations in the early 1980s. In recent years, Anderson *et al.* (1996) have tried to enhance the sulfur anomalous signal by collecting diffraction data using Cr K $\alpha$  X-

rays, while Kwiatkowski *et al.* (2000) constructed an X-ray target of gold, copper and chromium for pseudo-MAD experiments to be carried out at their home laboratory. Thus far, the anomalous signal of S or other atoms collected with Cr  $K\alpha$  radiation has mainly been used to determine the positions of these elements as a further independent check for correctness of the models.

In addition to the phasing power of soft X-rays in macromolecular crystallography, some studies performed at synchrotrons (Helliwell, 2002; Polikarpov *et al.*, 1997; Teplyakov *et al.*, 1998) have shown that longer wavelengths can actually be beneficial in the probing of small samples. This is a direct result of the increase in the scattering of the X-rays by the electrons in the sample as the wavelength increases and the reduction in absorption as the sample size decreases. Polikarpov *et al.* (1997) have suggested that wavelength be tuned to the sample size in order to maximize the diffracted energy.

More recently, advances in X-ray generator, optic and detector technologies, powerful computational algorithms for data processing, phase determination and density-modification programs and cryogenic techniques have made the method of sulfur SAD phasing using Cr  $K\alpha$  radiation feasible for routine structural determination in a home laboratory. In this paper, we describe a home-laboratory system designed specifically for chromium radiation and demonstrate that the data sets collected using this system are sufficient for sulfur SAD phasing. The system consists of an RU-300 rotating-anode X-ray generator with a Cr anode, Confocal MaxFlux (Cr CMF) multilayer optics specially designed for Cr  $K\alpha$  radiation by Osmic Inc., a square-pyramidal helium beam path and a large-aperture R-AXIS IV detector. The enabling technology of this experimental setup is the Cr CMF optic. It produces Cr  $K\alpha$  radiation with an intensity comparable to that of an RU-H2R/Blue optic (Yang *et al.*, 1999) at better than 94% spectral purity. The square-pyramidal He path allows the front window to be as close to the sample as possible in order to minimize the travel distance through air for the diffracted X-rays. The longer wavelength of Cr  $K\alpha$  radiation spreads out reciprocal space compared with Cu  $K\alpha$  radiation, so the  $300 \times 300$  mm aperture of the R-AXIS IV enables data collection to a reasonable resolution limit for phasing. In addition, the very small air gap between the imaging plate and front window of the detector make this detector more suitable for this experiment.

In addition to the enhancements in the hardware, improvements in software for data processing (Otwinowski, 1993; Otwinowski & Minor, 1997; Pflugrath, 1999), the new direct-methods algorithm in *SHELXD* (Sheldrick *et al.*, 2001) and maximum-likelihood-based heavy-atom refinement and phasing method in *SHARP* (de La Fortelle & Bricogne, 1997) were all essential to exploit the small anomalous signals from the weak anomalous scatterers such as sulfur and calcium.

Thaumatin and trypsin crystals were used in this experiment. In contrast to the previously published work on sulfur SAD phasing, data with much lower redundancy was needed to locate the positions of S atoms at low resolution ( $\sim 3.5$  Å)

using direct methods. The anomalous signal from the data provided sufficient phasing power to phase the protein data.

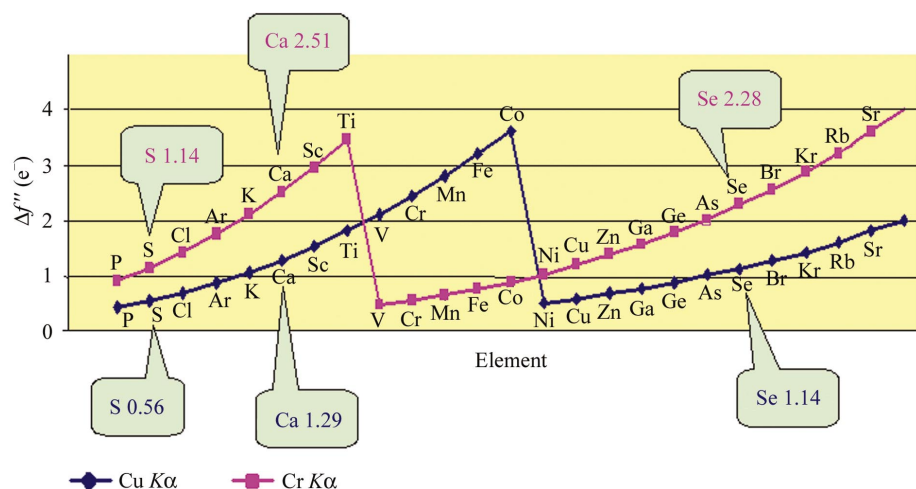
The results from these experiments reveal that sulfur SAD phasing with Cr  $K\alpha$  radiation is possible for routine protein-structure determination. Cr radiation can also enhance the anomalous signal of not only sulfur but also other intrinsic elements such as Ca, K and Zn and many heavy atoms commonly used in protein crystallography such as Hg, Pt, Xe and Br (Fig. 1). Many cases of phasing can benefit from data collected with Cr radiation.

## 2. Materials and methods

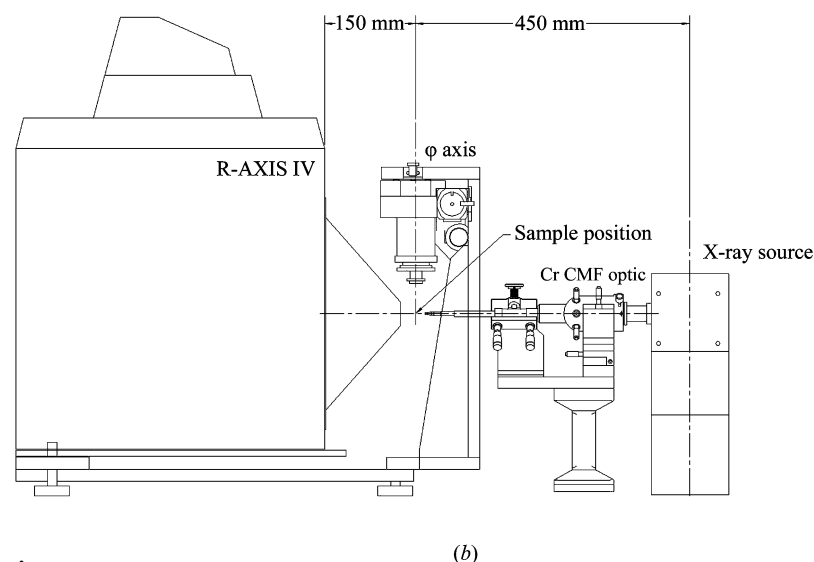
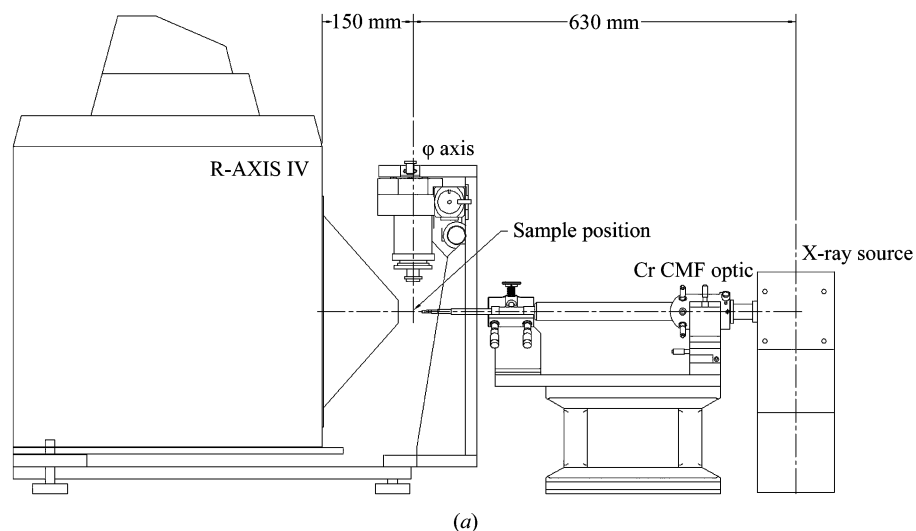
### 2.1. Experimental setup and optimization

As discussed in §1, the primary advantage of chromium radiation was recognized a long time ago. Some explorative experiments were conducted to phase protein data in the 1980s and 1990s. However, its application is still much less than might be expected. We believe this is because of the lack of intense monochromatic Cr radiation from a home source. The possibility of using Cr radiation to measure weak anomalous scattering signals became a reality with the new technology of multilayer optics. The Osmic CMF optic (Cr CMF) in the experimental setup enables the success of the experiments reported here. The optic consists of two graded multilayer surfaces in a confocal configuration (Fig. 2). The curvature and  $d$ -spacing of each surface were optimized for chromium radiation. The distance between the source and the focus point is designed to be 650 mm. With the fixed focus distance and the different apertures in the collimator, the configuration of the data-collection system can be optimized in several different ways. Three configurations were evaluated to determine the best conditions for collecting the anomalous signal from S atoms. Configuration *A* has the focal point of the optic at or near the sample position in order to maximize the flux at the crystal with a 0.5 mm aperture collimator. Configuration *B* has the focal point of the optic on or near the detector plane with a 0.5 mm aperture collimator to achieve the smallest reflection size. Configuration *C* keeps the focal point on the detector but uses a 1.0 mm aperture collimator. The criteria used in the evaluation were the flux and divergence,  $\langle I/\sigma(I) \rangle$  and sulfur peak heights in the anomalous difference Fourier map for data collected on the same thaumatin crystal.

The data in Table 1 show that configuration *C* produces the highest sulfur peaks in the anomalous difference Fourier map. As expected, configuration *A* maximizes the flux on the sample, but the diffracted X-rays start to diverge after the focal point. The full-width at half-maximum (FWHM) of the smallest reflections is 0.58 mm in both the  $X$  and  $Y$  direction on the detector at a 150 mm sample-to-detector distance. This contrasts with reflections that are 25% smaller in configuration *C*. The higher divergence makes the spots at high resolution larger by 30% than either configuration *B* or *C*. The result is that the  $\langle I/\sigma(I) \rangle$  of configuration *A* is very close to configuration *C* in spite of the doubling of the flux compared



**Figure 1**  
A diagram showing  $\Delta f''$  of the elements from phosphorus to strontium for Cu  $K\alpha$  and Cr  $K\alpha$  radiation.



**Figure 2**  
Schematic diagram of the configurations of the chromium system used in the experiment. (a) The configuration *A* with a pinhole size of 0.5 mm; (b) the configurations *B* and *C* with pinhole sizes of 0.5 and 1.0 mm, respectively.

with that of configuration *C*. Higher flux and longer wavelength radiation without increasing  $\langle I/\sigma(I) \rangle$  does little to enhance the anomalous scattering signal but can cause greater radiation damage in the crystal within a given time. Configuration *B* used a 0.5 mm aperture in the collimator. It has the smallest spot size of the reflections on the detector at 150 mm distance and low divergence, but the X-ray flux was sufficiently reduced compared with both *A* and *C* since the 0.5 mm aperture was too small for the divergence designed into this particular optic. The small aperture close to the optic has the net effect of selecting the more divergent rays, resulting in a somewhat higher overall divergence. The comparison between the results using the thaumatin data from configurations *B* and *C* further shows that configuration *C* is the best configuration of the three in terms of maximizing the sulfur anomalous signal, at least for this particular sample. Configuration *C* was used in the experiments reported in this paper.

The absorption of Cr radiation by materials in the beam path other than the crystal is one of the major factors that can reduce the signal of the reflections. With a standard home-laboratory system consisting of a chromium rotating-anode generator and an R-AXIS imaging-plate detector at a 150 mm sample-to-detector distance, the total absorption of the air and conventional black paper of the detector results in more than 50% loss of intensity at normal incidence. The transmission of the diffracted X-rays can be calculated using the equation

$$T(2\theta) = 1 - \exp[-(\mu_a D_a x_1 + \mu_p D_p x_2) / \cos(2\theta)], \quad (1)$$

where  $\mu_a$  and  $\mu_p$  are the mass attenuation coefficients of air and the window material, respectively,  $D_a$  and  $D_p$  are the densities of air and the window material, respectively,  $x_1$  is the air distance the photons travel from the sample to the detector and  $x_2$  is the thickness of the window. When the crystal-to-detector distance is 150 mm, less than half of the intensity of a

reflection can be recorded on the detector (Fig. 3).

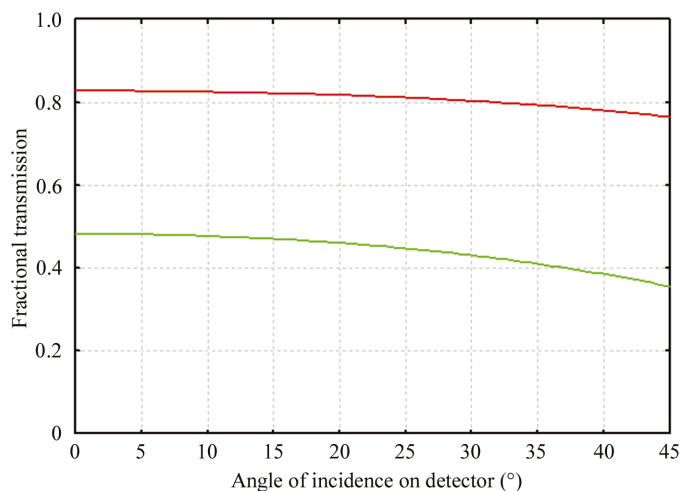
In order to reduce the loss of intensity of the diffracted X-rays, a helium path was constructed (Fig. 4). The small front window allows the He path a close approach to the back of the beam stop and minimizes the air path for all reflections. The materials of the front and back window are Mylar and carbon-filled polymer, respectively. The back window of the beam path is the front window of the detector, providing an air- and light-tight seal. This configuration reduces the loss by absorption to less than 20%. The width of the support arm of the beam stop was also trimmed to further shorten the air-travel distance for all the reflections.

As mentioned previously, an R-AXIS IV was used as the detector. The large aperture of the R-AXIS IV allows data collection to a reasonable resolution for phasing ( $2.7 \text{ \AA}$  for a 150 mm sample-to-detector distance). This resolution is limited by the helium beam path we built and not by the detector itself.

The chromium X-ray generation system used in the experiments consists of a Rigaku RU-300 rotating-anode generator running at 4.5 kW equipped with the Cr CMF multilayer optics described above.

## 2.2. The correlation between the sulfur anomalous signal and data parameters

A test crystal of *Thaumatooccus deniellii* thaumatin was used. Thaumatin purchased from Sigma was crystallized according to the conditions documented in the literature with slight modifications (Van der Wel *et al.*, 1975). The crystal was mounted in an arbitrary orientation and data were collected at cryogenic temperatures in a single continuous scan of  $585^\circ$ . The data were divided into 13 independent  $45^\circ$  scans and processed using *HKL2000* (Otwinowski, 1993; Otwinowski & Minor, 1997). Every  $45^\circ$  data set was treated equally using the



**Figure 3**  
Plot of transmission of Cr  $K\alpha$  radiation versus the angle of incidence on the detector. The green line is the transmission curve when the crystal-to-detector distance is 150 mm and no helium path is installed. The red line is the transmission curve when the crystal-to-detector distance is 150 mm and the helium path (Fig. 4) is installed. The values are relative to no loss of photons.

**Table 1**

Comparison of the different optic configurations.

Configuration	A	B	C
Source-to-sample distance (mm)	630	450	450
Collimator (mm)	0.5	0.5	1.0
Flux† (PIN diode units)	10.8	3.5	5.5
FWHM $X^\ddagger$ (mm)	0.58	0.40	0.44
FWHM $Y$ (mm)	0.58	0.33	0.40
Divergence $X$ (mrad)	2.2	1.6	1.1
Divergence $Y$ (mrad)	2.6	1.7	1.0
$R_{\text{merge}}^\S$ (%)	0.40	0.46	0.39
$\langle I/\sigma(I) \rangle$	13.5	8.4	12.2
Height of first sulfur peak¶	5.63	6.45	7.79
Height of second sulfur peak¶	4.51	4.32	5.16
Height of third sulfur peak¶	3.88	4.07	4.34

† The flux was measured at the output of the collimator with a Rigaku/MSD PIN diode detector. ‡ Full-width at half-maximum of direct-beam spot at 150 mm crystal-to-detector distance. § The same thaumatin crystal was used for the data collections with all three configurations. The collection sequence followed the alphabetic order.  $90^\circ$  of equivalent data was used in data processing and calculation of the sulfur peak heights. ¶ The three highest sulfur peaks in the anomalous difference Fourier map. The same thaumatin structural model was used to calculate the phases after rigid-body refinement.

*SCALEPACK2MTZ* and *TRUNCATE* (French & Wilson, 1978) programs to calculate the amplitude of structure factors ( $|F|$ ) and anomalous differences ( $\Delta F$ ) from the measured intensities ( $I$ ). The model (PDB code 1thw) was used as the initial model. The same cycles of rigid-body, overall and individual  $B$ -factor refinements were carried out to refine the same initial model against the different  $45^\circ$  data sets. The final refined models were used to calculate the phases ( $\varphi_{\text{calc}}$ ) using the *SFALL* program (Collaborative Computational Project, Number 4, 1994). Anomalous difference Fourier maps ( $|F^+| - |F^-|$ ,  $\varphi_{\text{CALC}} - 90^\circ$ ) were generated using the *FFT*



**Figure 4**  
The helium beam-path assembly, beam stop and Osmic Cr CMF optic.

**Table 2**  
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Thaumatococcus	Trypsin
Radiation (Å)	2.29	2.29
XG power (kW)	4.5	4.5
Detector and optics	R-AXIS IV, Osmic Cr CMF optic	R-AXIS IV, Osmic Cr CMF optic
Unit-cell parameters (Å, °)	$a = 58.2, b = 58.2, c = 150.7,$ $\alpha = 90.0, \beta = 90.0, \gamma = 90.0$	$a = 54.1, b = 58.3, c = 66.6,$ $\alpha = 90.0, \beta = 90.0, \gamma = 90.0$
Space group	$P4_12_12$	$P2_12_12_1$
Rotation range (°)	45	180
Resolution (Å)	2.7	2.7
$R(I)_{\text{merge}}$ (%)	4.0 (13.2)	3.6 (16.6)
Completeness (%)	93.9 (58.6)	89.4 (40.3)
Redundancy	2.9 (1.5)	3.2 (1.4)
$\langle I/\sigma(I) \rangle$	7.5 (2.3)	10.3 (2.8)
Data processing	HKL2000	d*TREK
No. atoms per AU	1551 (207 amino acids)	1765 (245 amino acids)
No. anomalous scatterers per AU	17 S	1 Ca <sup>2+</sup> and 14 S
Resolution range of data used for phasing (Å)	15–3.0	15.0–2.8
Experimental $\langle \Delta F \rangle / \langle F \rangle$ (%)	2.5	3.3
Calculated $\langle \Delta F \rangle / \langle F \rangle$ (%)	2.7	3.0

program (Ten Eyck, 1973). *PEAKMAX* (from the *CCP4* suite) was employed to list the peak heights of S atoms in the anomalous difference Fourier map. For all calculations described here, diffraction data within the resolution range 15.0–2.8 Å were used. This study was conducted to investigate the possible correlation of the anomalous signal of S atoms (the peak heights of S atoms in the anomalous difference Fourier map) with the overall inverse scale factors, completeness of each data set and isotropic displacement parameters of the S atoms in the final refined structures.

### 2.3. Sulfur SAD phasing

Two test proteins were used in this experiment: bovine pancreatic trypsin and *T. deniellii* thaumatococcus. Thaumatococcus was crystallized in the same way as mentioned above. Trypsin was purchased from Sigma and crystallized using the procedure described in the literature (Bode & Schwager, 1975).

As described above, the crystals were mounted in arbitrary orientations and collected at cryogenic temperatures in a single continuous scan. Two screening images at 0 and 90° were collected for assessment. The unit cell and orientation matrix of the crystals were determined with the reflections from the two images. A data-collection strategy was calculated with *d\*TREK* (Pflugrath, 1999) in order to determine the best rotation range to collect complete data (~99%). The first 45° of thaumatococcus data and 180° of trypsin data appear to be sufficient to phase the entire structures of both proteins. The data-collection statistics are listed in Table 2.

The thaumatococcus and trypsin data were treated with *SCALEPACK2MTZ* or *DTREK2MTZ* and *TRUNCATE* to calculate the amplitude of the structure factors ( $|F|$ ) and anomalous differences ( $\Delta F$ ) from the measured intensities ( $I$ ).

The initial positions of S atoms in the thaumatococcus crystal were located using the direct-methods program *RANTAN* (from the *CCP4* suite). The coordinates of the first three peaks found by *RANTAN* were transferred to the phase-refinement program *SHARP* for refinement and to calculate the initial phases. The solvent-flattening procedures *DM* and *SOLOMON* were used to resolve the phase ambiguity. The residual and anomalous difference Fourier maps produced by *SHARP* were examined in order to locate further sulfur peaks. The new sulfur sites were included in subsequent cycles of phase refinement and calculation using *SHARP*. The above cycle was repeated a few times in order to locate all the sulfur positions (Table 4).

The direct-methods program *SHELXD* was employed to locate the positions of the Ca<sup>2+</sup> ion and S atoms for the trypsin data. The anomalous difference data were prepared using *SHELXPRO* (from the *SHELX97* suite). The positions of the Ca<sup>2+</sup> and all the S atoms were found in the first run of *SHELXD*. The coordinates of the Ca<sup>2+</sup> and S atoms were input into *SHARP* for refinement and phase calculation, and phases were calculated at a resolution of 3.5 Å. Solvent flattening was used to resolve the phase ambiguity. An anomalous difference Fourier map was calculated using the SAD phases of the Ca<sup>2+</sup> and S atoms at a resolution of 2.8 Å. At this resolution, the disulfide bonds were resolved by visual recognition of the sulfur-peak shapes. The position of each S super-atom used in the refinement and phase calculation was replaced with the positions of two S atoms as described previously by Yang & Pflugrath (2001). Electron-density maps at the three different resolutions 3.5, 3.0 and 2.8 Å were created using the new SAD phases after solvent flattening.

The improvements in the electron-density maps were visualized using the program *O* (Jones *et al.*, 1991). The map correlation coefficient of the  $F_o$  map was calculated by *OVERLAPMAP* (Jones & Stuart, 1991) with the  $2F_o - F_c$  map ( $F_c$  from the final atomic model) used as the reference map.

## 3. Results and discussion

### 3.1. The correlation between the strength of the anomalous scattering of S atoms and data-collection parameters

Cryocrystallography of biological macromolecules has become *de rigueur* for data collection at both synchrotron beamlines and home laboratories. However, radiation damage is still observed and may restrict the amount of structural information that can be extracted from the samples. The work of Müller *et al.* (2002) indicates that radiation damage to

**Table 3**

Comparison of the sulfur-peak heights in the anomalous difference Fourier maps of the first and ninth 45° scans.

	First 45° scan	Ninth 45° scan
Completeness (%)	96.2	96.2
Overall scale factor†	1.00	1.01
$\langle I/\sigma(I) \rangle$ of the resolution shell 2.90–2.80 Å‡	30.1	27.3
Average peak heights ( $\sigma$ )	9.52	7.36
Peak height of Met112 ( $\sigma$ )	9.49	9.29
Average peak height of eight S super-atoms ( $\sigma$ )	9.53	7.12
Average sulfur isotropic displacement parameter (Å <sup>2</sup> )	12.4	16.9
Isotropic displacement parameter of Met112 (Å <sup>2</sup> )	14.9	19.9
Average isotropic displacement parameter of disulfide atoms (Å <sup>2</sup> )	12.2	16.7

† The overall scale factor was calculated within the resolution range 15.0–2.8 Å using the program *FHSCAL* (from the *CCP4* suite) using the first 45° scan as the reference scan. ‡ The  $\langle I/\sigma(I) \rangle$  of the resolution shell 2.90–2.80 Å was calculated with *SCALEPACK* using the same error model.

biological macromolecules may increase significantly above a wavelength of 1.3 Å using synchrotron radiation. The damage is expected to be random and to occur at all the amino-acid residues. This leads to a gradual degradation of the resolution of the diffraction pattern and to reduction of the diffraction signal. Recently, experimental evidence indicates that damage may be localized at specific sites within proteins (Burmeister, 2000; Ravelli & McSweeney, 2000; Weik *et al.*, 2000, 2001, 2002). The damage includes disulfide-bond cleavage and alterations of both cysteine and methionine residues. The Cr *K*α source does not have a flux comparable to that of synchrotrons; nevertheless, we observe similar damage to crystals because of the long wavelength. Both the general degradation and the damage to the S-atom bonds can seriously affect the strength of SAD phasing using sulfur anomalous scattering collected with Cr *K*α radiation. Therefore, we investigated the correlation between the sulfur anomalous signal and the data-collection parameters in order to reveal the optimal strategy of the data collection.

The thaumatin crystal used in this study was approximately 0.15 × 0.15 × 0.2 mm in size. A 585° data set was collected in a single scan and divided into 13 non-overlapping 45° scans. This long data collection allowed the first five 45° scans to be recollected after the initial 360° rotation. The data were integrated using *HKL2000* and imported into the *CCP4* suite. The sulfur-peak heights in the anomalous difference Fourier map were calculated with data in the resolution range 15–2.8 Å for each 45° scan. The corresponding completeness and inverse overall scale factor were also calculated using the program *FHSCAL* (from the *CCP4* suite). The isotropic displacement parameter (*B* factor) of each S atom was refined against each 45° data set.

Fig. 5 displays the average anomalous difference peak height of the S atoms, the completeness, the average isotropic displacement parameter for the S atoms and the inverse scale factor of each independent 45° scan across the entire 585° angular range.

The inverse scale-factor plot suggests that there was a general degradation in the data of about 5%. The average isotropic displacement parameter plot shows, assuming linear behavior, a 55% increase in the average isotropic displace-

ment parameters for the S atoms. The plot of completeness *versus* centroid of each 45° scan of data shows a systematic variation that resulted from the space group (*P*4<sub>1</sub>2<sub>1</sub>2), orientation of the crystal and the binning of the data into 45° scans. The variation appears to be sinusoidal, with a mean equal to the mean completeness.

The average anomalous difference Fourier peak heights of the S atoms in the map of each 45° of data gives the appearance of a damped sinusoidal curve with general decay. This is evident for at least the first six 45° scans. We have applied the following model to the

anomalous difference Fourier peak height data,

$$F(x) = a_{1a} + a_{2a} \sin(a_3x + a_4) \exp(-a_{5a}x) + a_{6a}x, \quad (2)$$

where  $a_{1a}$  represents the mean anomalous difference Fourier peak at the start of data collection,  $a_{2a}$  represents the variation in anomalous difference Fourier peak height as a result of the change in completeness for each 45° scan of data,  $a_3$  represents the frequency of the alternation in completeness, which by virtue of the binning into 45° scans is set to 4, and  $a_4$  represents the phase shift in the completeness, which has been set to 0. We believe that  $a_{5a}$  models the accelerated decay in the anomalous signal arising from chemical changes in the sites of the anomalous scatterers, *i.e.* S atoms, and that  $a_{6a}$  models the general decay of the sample from radiation damage. Note a simplified model to fit the completeness is used,

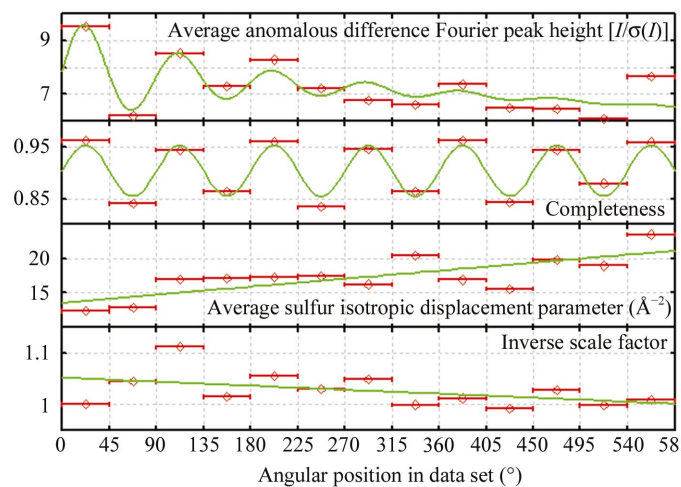
$$F(x) = a_{1c} + a_{2c} \sin(a_3x + a_4), \quad (3)$$

where the coefficients have the same meaning as above.

As shown in Fig. 5, after 270° of data have been collected the scans display much less sinusoidal behavior in the anomalous peak height and the average anomalous peak height tends to decrease as data collection proceeds regardless of the change in the completeness. This suggests that the S atoms may have suffered too much radiation damage relative to the change in signal from completeness. Furthermore, direct comparison of the data of the first 45° and ninth 45° scans (that is, of equivalent data sets differing only in the time of collection) also supports this hypothesis (Table 3). The ninth 45° scan was collected at the same crystal rotation range as the first 45° scan, but after 360° of data collection.

The first and ninth scans may be used to help deconvolute the combined effect of variation of the anomalous signal arising from completeness and from radiation damage in order to observe the effect of the radiation damage. The average anomalous peak height was reduced by ~23% after 360° of data collection. A comparison of inverse scale factors and  $\langle I/\sigma(I) \rangle$  in the last resolution shell (2.90–2.80 Å) indicate that a small level of overall decay of the crystal occurred during the data collection despite there being no visible reduction of resolution in the diffraction images. Such decay may possibly decrease the anomalous signals of the anomalous scatterers,

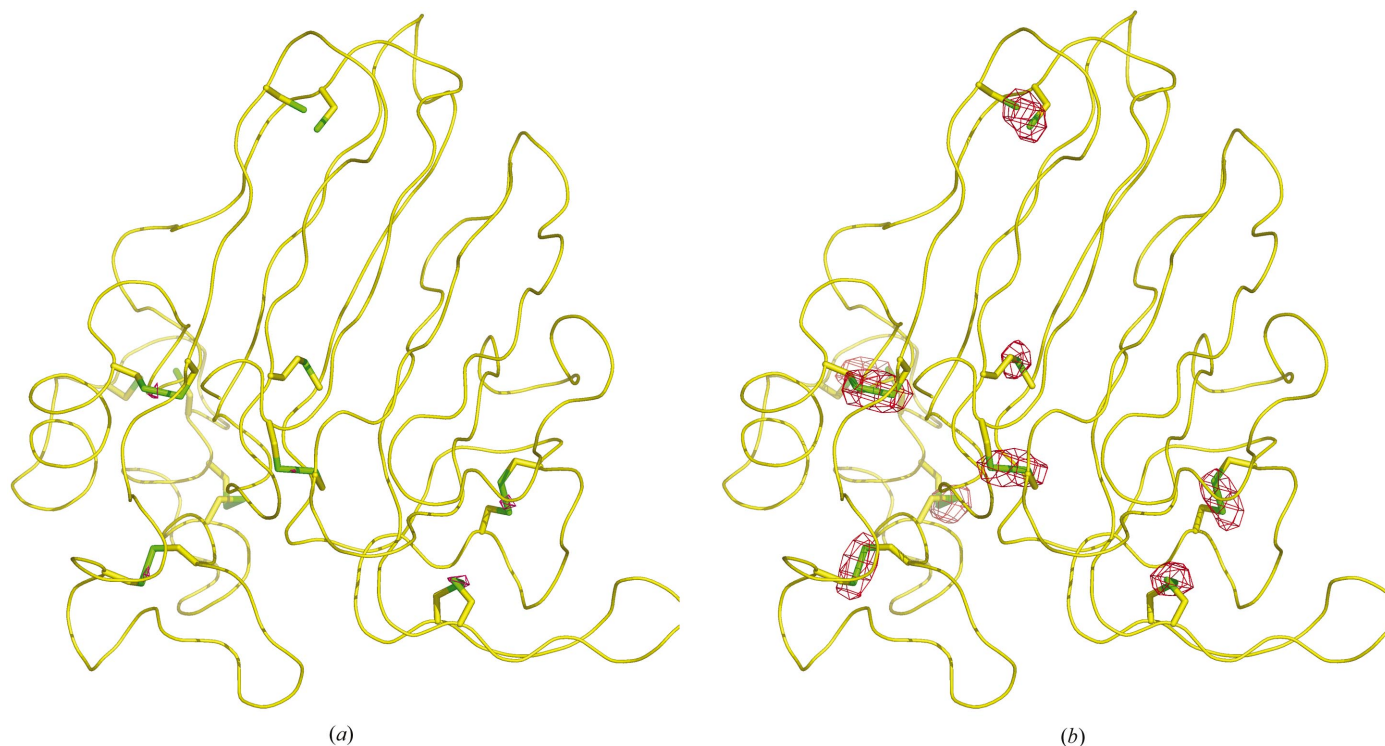
including S atoms. When the peak heights were assigned to each of the S atoms of the thaumatin molecule, it was observed that the anomalous peak height of the SD atom of Met112 was reduced to a much smaller extent ( $\sim 2\%$ ) compared with the  $\sim 25\%$  reduction of average peak height of the S super-atoms (disulfide linkages). Previous studies by others on radiation



**Figure 5**  
Plots of the average anomalous difference peak height of the S atoms, the completeness of the Bragg reflections, the average isotropic displacement parameter for the S atoms and the inverse scale factor for the 13 independent and consecutive  $45^\circ$  scans across the entire  $585^\circ$  rotation range. The green lines are the mathematic models to fit the data. For details, see §3.1.

damage have revealed that disulfide bonds have been found to be the most radiation-sensitive moieties in several proteins (Burmeister, 2000; Ravelli & McSweeney, 2000; Weik *et al.*, 2000). The immediate structural environment around the sulfurs, for example solvent accessibility, seems to be one of the parameters that affect the radiation-sensitivity of a disulfide bond. The structure of thaumatin illustrates that the Met112 residue is buried in the deep core of protein and that the cysteine residues involved in the eight disulfide bonds have a relatively higher solvent accessibility compared with that of Met112. The combination of the residue type and the surrounding environment of these residues may result in the different reduction of the anomalous signals of S atoms.

The isotropic displacement parameter of each S atom was refined using both data sets. The average isotropic displacement parameter increased by 36% from the first  $45^\circ$  scan to the ninth  $45^\circ$  scan. This increase further supports the hypothesis that radiation damage occurred during data collection. This also suggests that more radiation damage may occur at the S-atom sites than elsewhere in the crystal lattice. However, the isotropic displacement of the SD atom of Met112 has almost the same percentage increase ( $\sim 34\%$ ) as the SG atoms of cysteine residues ( $\sim 36\%$ ), despite the observation that its anomalous peak height is reduced by much less ( $\sim 2\%$ ) compared with the reduction of the average peak height of S super-atoms ( $\sim 23\%$ ). So far, the results of our experiments are not able to give any conclusive answer on the direct relationship between the isotropic displacement parameters and the anomalous signal because more complicated changes may be induced by radiation damage but the



**Figure 6**  
The anomalous difference Fourier maps of the same crystal of thaumatin contoured at  $4\sigma$ . (a) From data collected with Cu  $K\alpha$  radiation; note that only four peaks above  $4\sigma$  are barely visible. (b) From data collected with Cr  $K\alpha$  radiation; all the expected peaks are above  $4\sigma$ .



radiation damage in the crystal is reflected by both of them. In summary, both completeness and radiation damage can directly affect the observed anomalous signal. Higher data completeness collected as fast as possible seems to preserve more of the anomalous signal when using Cr  $K\alpha$  radiation. Thus, the strategy to collect the data quickly becomes more important for Cr  $K\alpha$  data collection than in the case of Cu  $K\alpha$  data collection. A strategy of merging short data sets of multiple fresh crystals may mitigate the systematic and random errors and increase the usable anomalous signal compared with a strategy of collecting data from a single crystal for the same total exposure time.

### 3.2. The strength of anomalous scattering of S atoms with Cr $K\alpha$ radiation

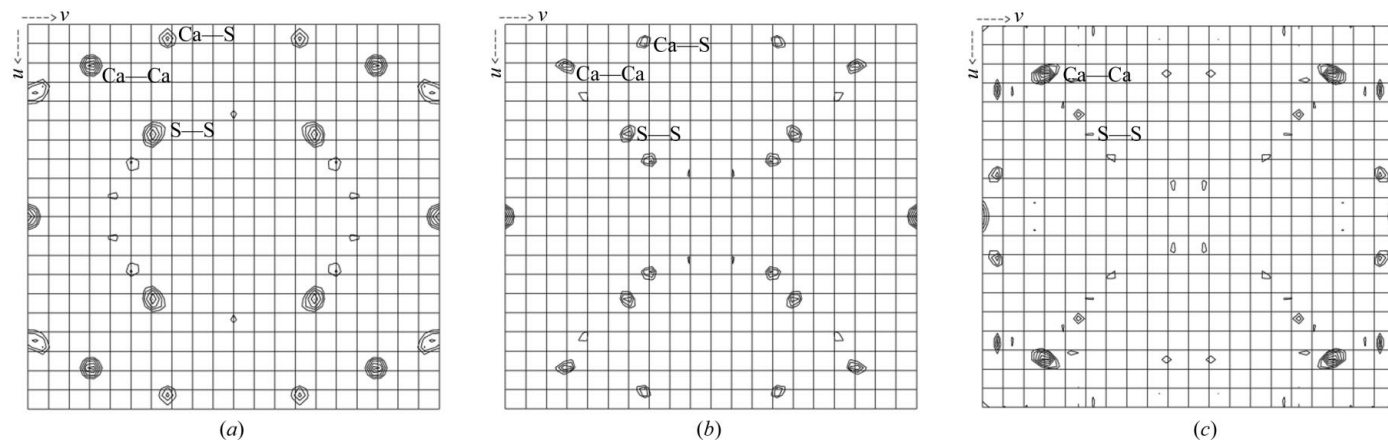
To demonstrate the increase in the strength of anomalous scattering using Cr  $K\alpha$  radiation, data from a fresh crystal of thaumatin were collected at 93 K using our standard conditions. Data were first collected with Cu  $K\alpha$  radiation and then with Cr  $K\alpha$  radiation. Based on the knowledge we gained above about the importance of completeness, we chose data-collection strategies that gave the  $45^\circ$  rotation range with the highest completeness for both the Cu  $K\alpha$  and Cr  $K\alpha$  data sets. The crystal-to-detector distances were kept the same (150 mm) for both data collections. The exposure times were adjusted to so that the crystal received an equivalent dose. Both data sets were processed using *HKL2000*. The anomalous difference Fourier maps were generated from the processed data within the resolution range 15–2.8 Å (Fig. 6). The same structural model of thaumatin was used in the rigid-body refinement against both data sets. The refined models were used to calculate the phase angles.

In the map calculated from the Cu  $K\alpha$  data only two anomalous peaks are higher than  $4\sigma$ . In the map calculated from the Cr  $K\alpha$  data nine peaks (one for the SD atom of Met112 and eight for the S super-atoms of the disulfides) are

above  $4\sigma$ . This indicates that the observable anomalous signal of S atoms is greatly increased under similar experimental conditions when Cr  $K\alpha$  radiation is used.

Our previous experience shows that finding the anomalous scatterer sites is the most difficult step in some cases of SAD phasing. An experiment was designed to investigate the possible improvement of the signal from S atoms observed in an anomalous difference Patterson map using Cr  $K\alpha$  radiation. Since the sulfur–sulfur vectors are rarely observed on the Patterson map with Cu  $K\alpha$  radiation, a  $1100^\circ$  rotation data set was collected on a fresh trypsin crystal using Cu  $K\alpha$  radiation first. A second data set of  $360^\circ$  was then collected on the same crystal using Cr  $K\alpha$  radiation. The parameters of the data collections were set up in the same way as the data collection of thaumatin in order to obtain comparable data. The data sets were processed using *d\*TREK*. The anomalous difference Patterson maps were calculated with processed data by increasing the total rotation range in  $90^\circ$  steps.

Fig. 7 shows the Harker sections in which the sulfur–sulfur peak appears. The sulfur–sulfur peak is observed in the Harker section of only the full  $360^\circ$  Cr  $K\alpha$  data. The same sulfur–sulfur peak did not appear until  $1020^\circ$  of copper data were included. This peak is so weak that it would be hard to distinguish above the noise without prior knowledge. The calcium–sulfur peak also shows the difference in the strength of the anomalous signals between Cr  $K\alpha$  and Cu  $K\alpha$  radiation. The calcium–sulfur cross-vectors display a strong peak on the Harker section of the  $360^\circ$  of Cr  $K\alpha$  data. The same peak is not observed on the Harker section from  $1020^\circ$  of Cu  $K\alpha$  data. This simple comparison between the Harker sections again indicates that Cr radiation can provide a great enhancement to the signal observed from the anomalous scattering of S atoms. It also implies that finding the sites of S atoms using direct methods and anomalous difference Patterson techniques may be much easier and require much less data when using Cr  $K\alpha$  radiation. This hypothesis was tested and confirmed in the example cases described in the next section.



**Figure 7**

Harker sections ( $w = 0$ ) of the anomalous difference Patterson maps of trypsin. (a) Calculated anomalous difference Patterson map, (b) anomalous difference Patterson map using  $360^\circ$  of data collected with Cr  $K\alpha$  radiation, (c) anomalous difference Patterson map of  $1020^\circ$  of data collected with Cu  $K\alpha$  radiation. The sections are contoured in increments of  $1\sigma$  starting at  $2\sigma$ .

**Table 4**

 Peak heights of S and super-S (S–S) atoms obtained from *RANTAN* and *SHARP*.

Position	<i>RANTAN</i> † (σ)	<i>SHARP</i> ‡ (σ), cycle 1	<i>SHARP</i> (σ), cycle 2	<i>SHARP</i> (σ), cycle 3
1 0.4658 0.9607 0.0201 (Cys126–Cys177)	13.5	18.3	16.6	17.7
2 0.3932 0.6438 0.0242 (Cys121–Cys193)	13.4	16.2	15.2	15.9
3 0.4056 0.9830 0.0986 (Cys5–Cys204)	11.9	18.4	14.6	14.8
4 0.0972 0.9569 0.0674 (Cys71–Cys77)	8.8		4.9	13.4
5 0.9726 0.3025 0.0949 (Cys159–Cys164)	8.6	5.5	15.9	12.0
6 0.3985 0.1436 0.0114 (Cys134–Cys145)	8.4		5.5	14.8
7 0.3488 0.8808 0.0443 Met112	7.7		4.8	13.4
8 0.1630 0.0478 0.0944 (Cys56–Cys66)	6.6	4.7	14.9	16.9
9 0.0717 0.2197 0.0573 (Cys149–Cys156)	5.7	4.9	14.3	13.5
10 0.3729 0.7119 0.1021		4.7		
11 0.3515 0.2292 0.1021		4.3		

† The S and S super-atom peak heights output by *RANTAN* (direct methods) at a resolution of 3.5 Å. ‡ The anomalous peak heights of S and S super-atoms were found in the anomalous difference Fourier maps with *SHARP* at a resolution of 3.5 Å.

### 3.3. Phasing tetragonal thaumatin and orthorhombic trypsin structure from a single data set collected with Cr Kα radiation

*T. deniellii* thaumatin, with a molecular weight of 22 kDa, has 17 S atoms from one methionine and 16 cysteine residues. This is a sulfur-rich protein and provides a good example for the initial tests. In this experiment, a single data set of a native thaumatin crystal was collected on a 4.5 kW laboratory source at Cr Kα wavelength (2.29 Å) and its structure was phased from the intrinsic anomalous scattering of one S and eight super-S atoms (–SS–).

The thaumatin crystal was approximately 0.15 × 0.15 × 0.2 mm in size. It crystallized in the expected space group *P*<sub>4</sub><sub>2</sub><sub>1</sub><sub>2</sub>. During X-ray diffraction data collection, the crystal was maintained at cryogenic temperatures. The data-collection statistics are listed in Table 2.

The thaumatin data were collected and processed to the edge of the detector (2.7 Å), although the crystal diffracted beyond this resolution. The resolution of the diffraction data was limited by our initial choice of helium beam-path length, which resulted in a minimum crystal-to-detector distance of 150 mm. *d*\**TREK* was used to determine the optimal scan range to attain 99% completeness. Previous results (see above) showed that 45° of data were sufficient for phasing the thaumatin structure. Table 2 lists the data statistics of the 45° scan used in the phasing procedure. The overall redundancy was about 2.9-fold, with 93.4% overall completeness. The overall *R*<sub>merge</sub> is 4.0% and is 13.2% in the highest resolution (2.8–2.7 Å) shell. The overall  $\langle I/\sigma(I) \rangle$  for the averaged observations is 7.5. When compared with the data reported in other cases of sulfur SAD phasing (Dauter *et al.*, 1999; Yang & Pflugrath, 2001; Debreczeni *et al.*, 2003), there is much lower redundancy in these data. It should also be noted the overall  $\langle I/\sigma(I) \rangle$  is also lower. In general, sulfur SAD phasing is not considered to be possible with similar data sets collected with Cu Kα.

The enhancement of the anomalous signal with Cr Kα radiation can be shown using the metric  $\langle |\Delta F| \rangle / \langle F \rangle$ . According to Hendrickson & Teeter (1981), the expected average ratio of

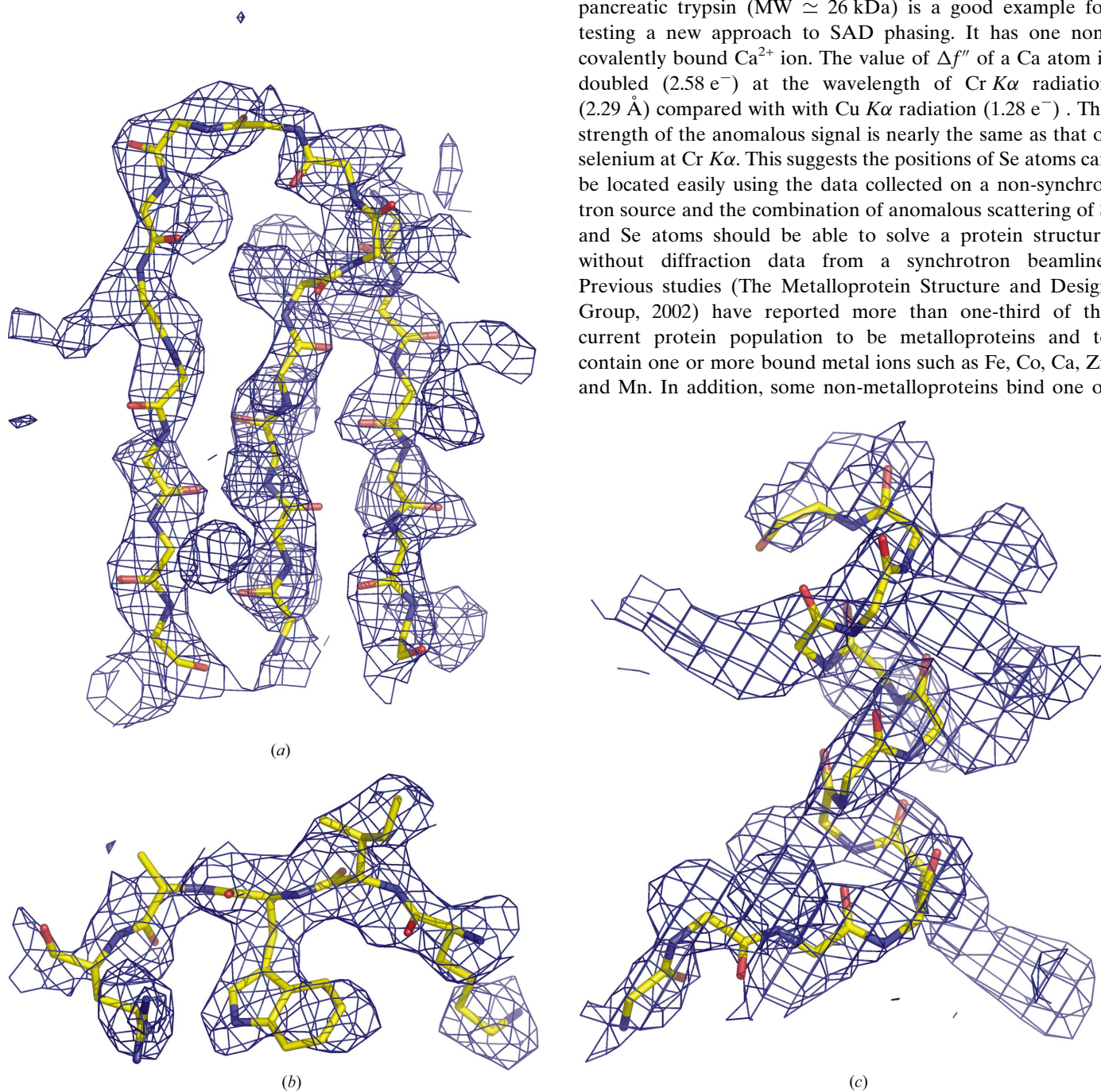
$\langle |\Delta F| \rangle / \langle F \rangle$  can be calculated as  $2^{1/2} (N_A^{1/2} \Delta f_A'' / (N_P^{1/2} Z_{\text{eff}}))$ . In the case of thaumatin, the number of anomalous scatterers (*N*<sub>A</sub>) in thaumatin is 17, the number of total non-H protein atoms (*N*<sub>P</sub>) is 1551 and the effective atomic number (*Z*<sub>eff</sub>) is ~6.7 for non-H protein atoms. 17 S atoms with  $\Delta f_A'' = 1.14 \text{ e}^-$  results in a value of 2.5% for the ratio of  $\langle |\Delta F| \rangle / \langle F \rangle$  at the wavelength of Cr Kα radiation, compared with a value of 1.2% for the same ratio with Cu Kα radiation. The overall experimental ratio of  $\langle |\Delta F| \rangle / \langle F \rangle$  is 2.7% to the resolution limit of 2.7 Å. The deviation from the expected value is because of the approximate nature of the Hendrickson and Teeter formula and experimental

error. The increase in the strength of the anomalous scattering signal with Cr Kα suggests that it is feasible to use fewer data for phasing the thaumatin structure by sulfur SAD than with Cu Kα and our results are consistent with this.

The direct-methods program *RANTAN* (from the *CCP4* suite) was used to locate the positions of the S atoms. Data within the resolution range 10.0–3.5 Å were used in the initial run. Table 4 shows the peak lists from *RANTAN* and *SHARP* and their corresponding residues.

The *RANTAN* result lists the positions of all the S atoms. Because the first three peaks were significantly higher than the others, they were input into *SHARP* as three S atoms to be refined and for which SAD phases would be calculated. A total of 3434 anomalous differences to a resolution of 3.5 Å were used. The figure of merit was 0.298. The solvent-flattening procedure of *DM* (Cowtan, 1994) was employed to resolve the phase ambiguity. The newly calculated SAD phases were used to calculate the anomalous difference Fourier map ( $|F^+| - |F^-|$ ,  $\varphi_{\text{SAD}} - 90^\circ$ ) and residual map using *SHARP*. Three more ‘anomalous’ peaks of significant height were found in the peak list. These three positions were added to the list for refinement in *SHARP* as S atoms. The next highest peak was 3.3σ (not shown in Table 4). The second cycle of refinement and phase calculation was carried out with these six sulfur positions. The six S-atom positions were refined and resulted in reasonable occupancies and increased peak heights. The figure of merit increased to 0.375. The same procedure of performing a *SHARP* run, calculating the anomalous difference Fourier map and residual map was used to locate more sulfur sites. Three additional sites corresponding to the S super-atoms Cys71–Cys77 and Cys134–Cys145 and the S atom of Met112 (Table 4) were found using the obvious decrease between position 7 and the highest noise peak (3.1σ). The third cycle of refinement and phase calculation was performed with the addition of these newly found sulfur sites. This resulted in reasonable occupancies and increased peak heights for every sulfur site. The figure of merit reached 0.485. The anomalous difference Fourier map and residual map were also calculated in order to locate more

sulfur sites. However, no further new sites were found. Since the data used here had a maximum of 3.5 Å resolution, the disulfide bond is found as a single S super-atom in the middle of the disulfide bond instead of two S atoms. Only nine positions can be found from thaumatin data at a resolution of 3.5 Å. The nine positions were again input to *SHARP* as nine super-atoms for refinement and phase calculations at a resolution of 3.0 Å. The figure of merit after solvent flattening (*DM*) increased to 0.680. The experimental electron-density maps are displayed in Fig. 8.



**Figure 8**

Experimental electron-density maps of the thaumatin structure generated from data to 3.0 Å resolution and sulfur SAD phases: (a) electron density of the  $\beta$ -sheet (Met112–Cys126, Asn198–Phe203); (b) electron density of Trp51; (c) electron density of the helix (Gly165–Asp179). Only main-chain atoms are displayed for (a) and (c).

The electron density is well defined in the map at a resolution of 3.0 Å. The electron density shown in Fig. 8(a) is easily recognized as a  $\beta$ -sheet. Residue Trp51 (Fig. 8b) is also easily recognized. Some of the side chains of the C-terminal helix appear in Fig. 8(c). The map correlation coefficient of the  $F_o$  map calculated by *OVERLAPMAP* is 53% when the  $2F_o - F_c$  map ( $F_c$  from the final atomic model) was used as the reference map. The main chain of 207 amino acids was traced within a few hours using manual methods.

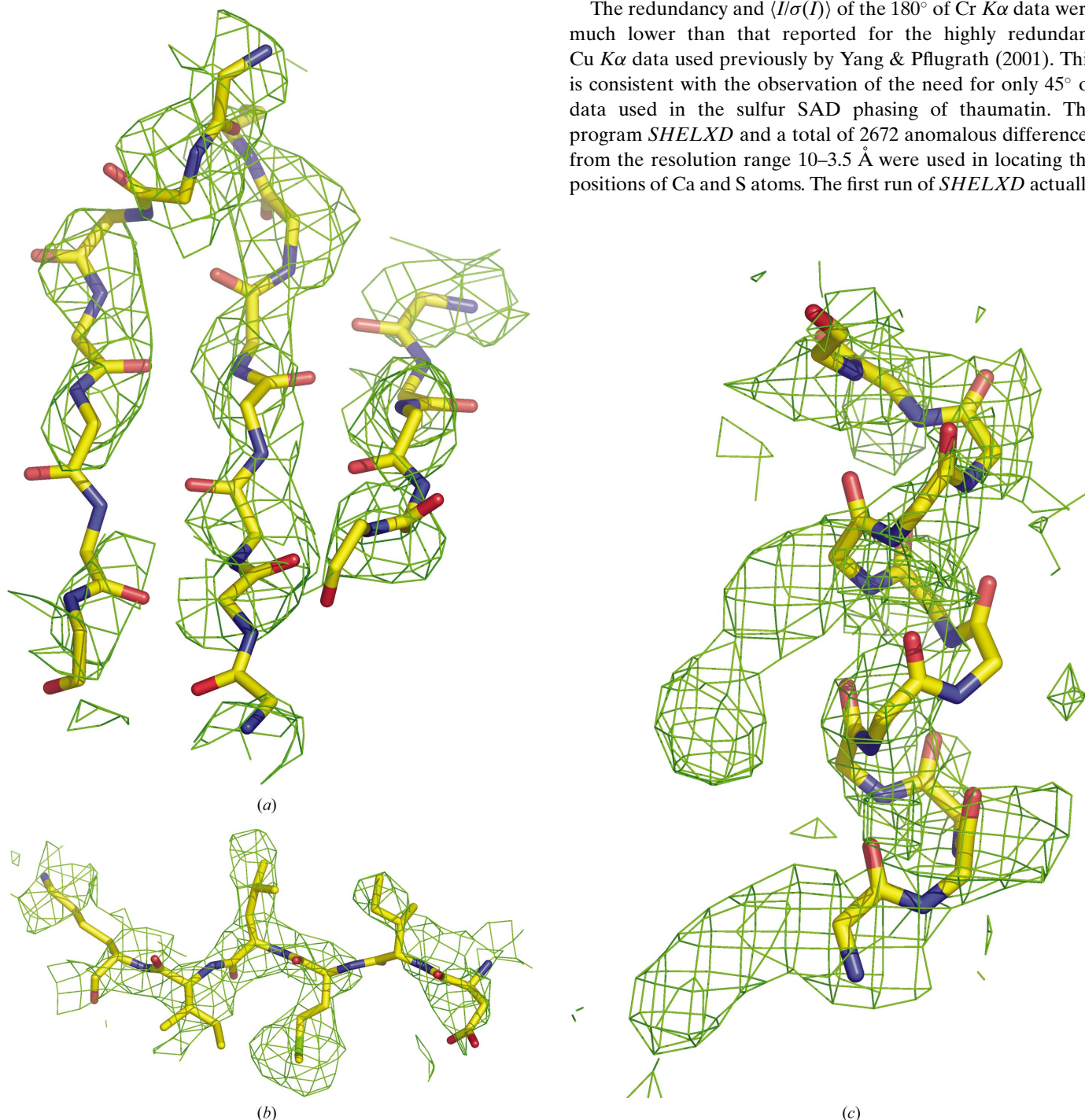
As reported previously by Yang & Pflugrath (2001), bovine pancreatic trypsin (MW  $\approx$  26 kDa) is a good example for testing a new approach to SAD phasing. It has one non-covalently bound  $\text{Ca}^{2+}$  ion. The value of  $\Delta f''$  of a Ca atom is doubled ( $2.58 e^-$ ) at the wavelength of Cr  $K\alpha$  radiation (2.29 Å) compared with with Cu  $K\alpha$  radiation ( $1.28 e^-$ ). The strength of the anomalous signal is nearly the same as that of selenium at Cr  $K\alpha$ . This suggests the positions of Se atoms can be located easily using the data collected on a non-synchrotron source and the combination of anomalous scattering of S and Se atoms should be able to solve a protein structure without diffraction data from a synchrotron beamline. Previous studies (The Metalloprotein Structure and Design Group, 2002) have reported more than one-third of the current protein population to be metalloproteins and to contain one or more bound metal ions such as Fe, Co, Ca, Zn and Mn. In addition, some non-metalloproteins bind one or

more ions such as K and Cl during crystallization. SAD phasing using Ca and S atoms with Cr  $K\alpha$  radiation reveals another possible way to efficiently phase metalloprotein diffraction data.

Trypsin was crystallized in the space group  $P2_12_12_1$ . A crystal  $0.1 \times 0.1 \times 0.2$  mm in size was selected for data collection. Although the crystal diffracted well beyond  $2.7 \text{ \AA}$ , the data were collected and processed to  $2.7 \text{ \AA}$  using  $d^*TREK$

for the same reason described above for thaumatin.  $d^*TREK$  was used to find the optimal rotation range for the data with 99% completeness. The data collection was carried out under cryogenic conditions and no noticeable decay was observed. The SAD phasing results showed a  $180^\circ$  scan of this trypsin data was sufficient for the program *SHELXD* to locate the Ca and S positions and subsequently phase the entire trypsin structure by SAD with *SHARP*. The statistics of the  $180^\circ$  scan of data are listed in Table 2.

The redundancy and  $\langle I/\sigma(I) \rangle$  of the  $180^\circ$  of Cr  $K\alpha$  data were much lower than that reported for the highly redundant Cu  $K\alpha$  data used previously by Yang & Pflugrath (2001). This is consistent with the observation of the need for only  $45^\circ$  of data used in the sulfur SAD phasing of thaumatin. The program *SHELXD* and a total of 2672 anomalous differences from the resolution range  $10\text{--}3.5 \text{ \AA}$  were used in locating the positions of Ca and S atoms. The first run of *SHELXD* actually



**Figure 9**

Experimental electron-density map of trypsin calculated using anomalous differences to  $3.5 \text{ \AA}$  and the SAD phases from the  $\text{Ca}^{2+}$ , six S super-atoms and two S atoms. (a) The electron-density maps of the  $\beta$ -sheet (Gln50–Ser54, Ser84–Val90, Ile103–Lys109); (b) the electron-density map of residues Asp102–Lys107; (c) electron density of the helix (Tyr234–Asn245). Only main-chain atoms are displayed for (a) and (c).

**Table 5**  
Peak list output by *SHELXD* from the 180° scan of trypsin data.

Peak	Fractional coordinates			Height
1	0.1949	0.5745	0.0557	99.9
2	0.5725	0.5277	0.2292	76.0
3	0.4648	0.7642	0.1054	65.3
4	0.3658	0.5222	0.3435	60.8
5	0.3429	0.7657	0.0535	57.6
6	0.3245	0.7597	0.3066	49.7
7	0.7615	0.4868	0.2317	49.1
8	0.3946	0.8503	0.3797	43.9
9	0.6349	0.7636	0.2537	41.8
10	0.1718	0.7658	0.0297	26.5
11	0.5162	0.6535	0.0694	20.7
12	0.0755	0.4541	-0.0014	18.8

gave the location of the  $\text{Ca}^{2+}$  ion, six S super-atoms and two S atoms (Table 5).

The first nine positions were input to *SHARP* for refinement and phase calculation, where the position of the highest peak was treated as the calcium site and the rest were treated as S atoms. Refinement was carried out at a resolution of 3.5 Å. The figure of merit is 0.45 to a resolution of 3.5 Å. The solvent-flattening procedure of *DM* was used to resolve the phase ambiguities and improve the electron-density maps. The resultant experimental electron-density maps are shown in Fig. 9.



**Figure 10**  
The 2.8 Å resolution anomalous difference Fourier map superimposed with the trypsin model. The map is contoured at  $3.5\sigma$ . The magenta sphere represents the  $\text{Ca}^{2+}$  ion. Two spherical densities cover the SD atoms of Met104 and Met180. Six elliptical densities mark the six disulfide bonds.

Although the electron-density map in Fig. 9 shows the recognizable secondary structures for trypsin structure, some discontinuous regions were seen in the map. Therefore, the refinement and phase calculation were extended to a resolution of 2.8 Å. The anomalous difference Fourier map was created and displayed in *O* (Fig. 10). As can be seen, the densities of the disulfide bonds and S atoms start to become distinguishable in the map at a resolution of 2.8 Å, with disulfide bonds often appearing as ellipsoids. The densities of the SD atoms of methionine residues appear to be spherical. Such differences become even more pronounced when higher resolution data is used. With the known bond length of the disulfide bond, the six S super-atoms were each replaced with two S atoms by identifying the boundaries of these densities.

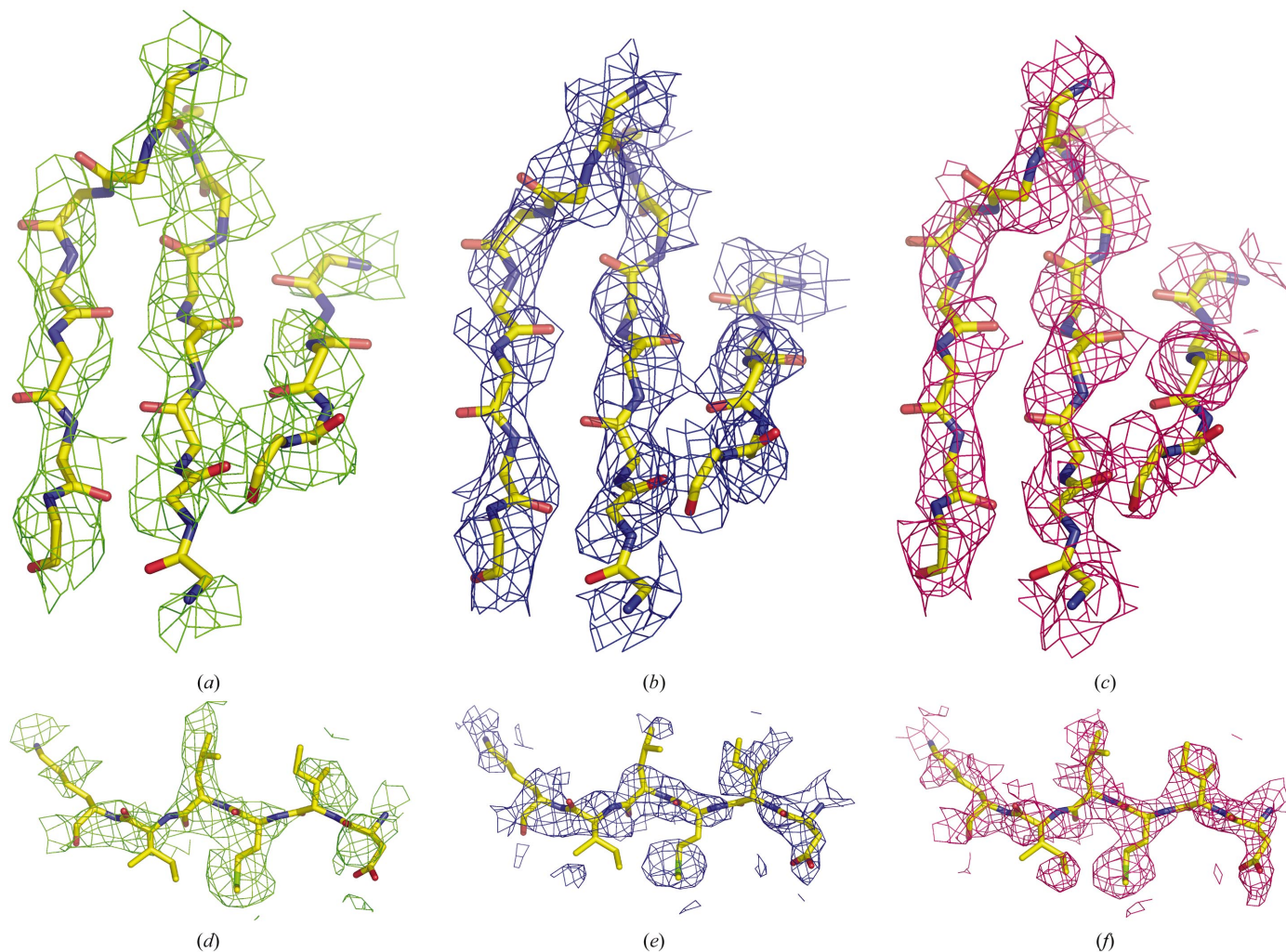
The positions of one  $\text{Ca}^{2+}$  and the 14 S atoms were input into *SHARP* for another cycle of refinement and phase calculation, which were carried out at three different resolutions: 3.5, 3.0 and 2.8 Å. The values of the figure of merit are 0.532, 0.465 and 0.432, respectively. The map correlation coefficients were also calculated from the  $2F_o - F_c$  map with the final refined structure as the reference map. These values are 0.391, 0.450 and 0.493, respectively. The experimental electron-density maps are displayed in Fig. 11.

Visual inspection of Fig. 11 clearly shows the electron-density maps were improved both by resolving the —SS— of a disulfide bond into two individual S atoms and by including data to higher resolution. This latter observation is consistent with the findings of Hendrickson & Teeter (1981) and our previous experience (Yang & Pflugrath, 2001). The 2.8 Å resolution electron-density map displayed well defined regions for almost all the secondary-structural features and can be traced both manually and automatically.

#### 4. Conclusion

High-throughput crystallography requires a method by which the structures of native proteins can be determined quickly and easily. In recent years, the combination of the MAD method and a tunable synchrotron radiation has become extremely powerful and popular in the field of structural biology for *de novo* structure solution, but difficulties still arise under certain circumstances. For instance, sometimes selenomethionine protein cannot be expressed in the system, crystals of selenomethionine protein are hard to grow or the crystals are too fragile to be handled. Other methods, such as MIR, SIR or SIRAS are also problematic owing to the trial-and-error nature in the preparation of heavy-atom derivatives.

In this regard, SAD phasing becomes increasingly attractive, especially SAD phasing using intrinsic anomalous scatterers such as sulfur. This work illustrates the advantages offered by sulfur SAD phasing, which has only recently become easily realised owing to several technological advances. These advances encompass several areas including a newly designed confocal multilayer optic that can produce Cr  $K\alpha$  radiation with sufficient brilliance and high spectral purity, improvements in beam optimization that reduce the loss of signals to a reasonable level, modern large-aperture



**Figure 11**

Electron-density maps of the trypsin structure calculated at 3.5 Å (green), 3.0 Å (blue) and 2.8 Å (magenta) resolution and SAD phases from one  $\text{Ca}^{2+}$  and 14 S atoms contoured at  $1\sigma$ . (a), (b) and (c) The electron density maps of  $\beta$ -sheet (Gln50–Ser54, Ser84–Val90 and Ile103–Lys109). Only main-chain atoms are displayed. (d), (e) and (f) The electron-density map covering residues Asp102–Lys107.

detectors that capture the broad diffraction pattern created by longer wavelength chromium radiation and advanced algorithms implemented in modern software that estimate reflection intensities with sufficient accuracy. The combination of all of these makes sulfur SAD phasing routinely achievable in the home laboratory for more samples than are achievable with  $\text{Cu K}\alpha$ .

The data-collection strategy appears to be important for data collection with Cr  $K\alpha$  radiation. A quickly collected and nearly complete data set retains most of the anomalous signal considering the combined yet inverse effects of completeness and radiation damage. In this case, the minimum data were used to solve the structure of thaumatin and trypsin; nevertheless, higher redundancy, resolution and  $\langle I/\sigma(I) \rangle$  can improve the accuracy of the measured Bijvoet differences and SAD phases and benefit phasing in many aspects.

We wish to thank B.-C. Wang and John Rose of the University of Georgia, and Gerard Bricogne of Global

Phasing Ltd for many useful discussions. We wish to thank Wladek Minor of the University of Virginia for assistance with *HKL2000*. We wish to thank Boris Verman and Licai Jiang of Osmic Inc. for the design and production of the Cr CMF optic. We wish to thank Akihito Yamano of Rigaku Corporation and Thad Niemeyer and Angela Criswell of Rigaku/MSI Inc for their assistance with this work.

## References

- Anderson, D. H., Weiss, M. S. & Eisenberg, D. (1996). *Acta Cryst.* **D52**, 469–480.
- Behrens, W., Otto, H., Stuhmann, H. B. & Heyn, M. P. (1998). *Biophys. J.* **75**, 255–263.
- Blow, D. E. (1958). *Proc. R. Soc. London Ser. A*, **247**, 302–336.
- Bode, W. & Schwager, P. (1975). *J. Mol. Biol.* **98**, 693–717.
- Borek, D., Minor, W. & Otwinowski, Z. (2002). Abstr. Am. Crystallogr. Assoc. Annu. Meet., Abstract P024.
- Broderson, D. E., de La Fortelle, E., Vornrhein, C., Bricogne, G., Nyborg, J. & Kjeldgaard, M. (2000). *Acta Cryst.* **D56**, 431–441.
- Burmeister, W. P. (2000). *Acta Cryst.* **D56**, 328–341.

- Chen, L. Q., Rose, J. P., Breslow, E., Yang, D., Chang, W. R., Furey, W. F. Jr, Sax, M. & Wang, B.-C. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 4240–4244.
- Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F. & Helliwell, J. R. (2001). *Acta Cryst. D57*, 1219–1229.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst. D50*, 760–763.
- Cowtan, K. (1994). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34–48.
- Dauter, Z. & Dauter, M. (1999). *J. Mol. Biol.* **289**, 93–101.
- Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
- Dauter, Z., Dauter, M. & Dodson, E. (2002). *Acta Cryst. D58*, 494–506.
- Dauter, Z., Dauter, M. & Rajashankar, K.R. (2000). *Acta Cryst. D56*, 232–237.
- Debreczeni, J. É., Bunkóczi, G., Ma, Q., Blaser, H. & Sheldrick, G. M. (2003). *Acta Cryst. D59*, 688–696.
- French, G. S. & Wilson, K. S. (1978). *Acta Cryst. A34*, 517.
- Helliwell, J. R. (2002). *J. Synchrotron. Rad.* **9**, 1–8.
- Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
- Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
- Jones, T. A., Zhou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst. A47*, 110–119.
- Jones, Y. & Stuart, D. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 39–48. Warrington: Daresbury Laboratory.
- Kwiatkowski, W., Noel, J. P. & Choe, S. (2000). *J. Appl. Cryst.* **33**, 876–881.
- La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
- Lemke, C. T., Smith, G. D. & Howell, P. L. (2002). *Acta Cryst. D58*, 2096–2101.
- Liu, Z.-J., Vysotski, E. S., Cheng, C. J., Rose, J. P., Lee, J. & Wang, B.-C. (2000). *Protein Sci.* **9**, 2085–2093.
- Müller, R., Weckert, E., Zellner, J. & Drakopoulous, M. (2002). *J. Synchrotron Rad.* **9**, 368–374.
- Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst. D57*, 996–1002.
- Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pflugrath, J. W. (1999). *Acta Cryst. D55*, 1718–1725.
- Polikarpov, I., Teplyakov, A. & Oliva, G. (1997). *Acta Cryst. D53*, 734–737.
- Ravelli, R. B. G. & McSweeney, S. M. (2000). *Struct. Fold. Des.* **8**, 315–328.
- Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, M. & Usón, I. (2001). *International Tables for Crystallography, Vol. F*, edited by M. G. Rossmann & E. Arnold, pp. 333–351. Dordrecht: Kluwer Academic Publishers.
- Stuhrmann, S., Bartels, K. S., Braunwarth, W., Doose, R., Dauvergne, F., Gabriel, A., Knöchel, A., Marmotti, M., Stuhrmann, H. B., Trame, C. & Lehmann, M. S. (1997). *J. Synchrotron Rad.* **4**, 298–310.
- Stuhrmann, S., Hütsch, M., Trame, C., Thomas, J. & Stuhrmann, H. B. (1995). *J. Synchrotron Rad.* **2**, 83–86.
- The Metalloprotein Structure and Design Group (2002). *Metalloprotein Database and Browser*, <http://www.scripps.edu/research/metallo>.
- Ten Eyck, L. F. (1973). *Acta Cryst. A29*, 183–191.
- Teplyakov, A., Oliva, G. & Polokarpov, I. (1998). *Acta Cryst. D54*, 610–614.
- Van der Wel, H., van Soest, T. C. & Royers, R. C. (1975). *FEBS Lett.* **56**, 316–317.
- Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.
- Weik, M., Bergès, J., Raves, M. L., Cros, P., McSweeney, S., Silman, I., Sussman, J. L., Houée-Levin, C. & Ravelli, R. B. G. (2002). *J. Synchrotron Rad.* **9**, 342–346.
- Weik, M., Ravelli, R. B. G., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 623–628.
- Weik, M., Ravelli, R. B., Silman, I., Sussman, J. L., Gros, P. & Kroon, J. (2001). *Protein Sci.* **10**, 1953–1961.
- Weiss, M. S., Sicker, T., Djinovic Carugo, K. & Hilgenfeld, R. (2001). *Acta Cryst. D57*, 689–695.
- Wu, C.-K., Dailey, H. A., Rose, J. P., Burden, A., Sellers, V. M. & Wang, B.-C. (2001). *Nature Struct. Biol.* **8**, 156–160.
- Yang, C., Courville, A. & Ferrara, J. D. (1999). *Acta Cryst. D55*, 1681–1689.
- Yang, C. & Pflugrath, J. W. (2001). *Acta Cryst. D57*, 1480–1490.